APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/596,112	05/31/2006	Nevenka Dimitrova	NL040163US	2183

24737   7590   05/22/2009
PHILIPS INTELLECTUAL PROPERTY & STANDARDS
P.O. BOX 3001
BRIARCLIFF MANOR, NY 10510

EXAMINER
TRUONG, DENNIS

ART UNIT	PAPER NUMBER
2169	

MAIL DATE	DELIVERY MODE
05/22/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

DETAILED ACTION

1.   This Action is responsive to Applicant's Amendments filed March 12, 2009.

Response to Amendment

2.   It is acknowledged that claims 1, 10-15, 17, 19, 23, 25-28 have been amended. And claims
30-32 were canceled in the previous amendment (March 27, 2008).

3.   Claims 1-8 and 10-29 are pending.

Response to Arguments

4.   Applicant's arguments with respect to the rejected claim have been fully considered but
are not persuasive.

5.   On page 10-11, Applicant argues that although Lamkin indicates that the intrinsic
content, extrinsic content and the processor are in the same device, and the apparatus is able to
go to a specific location within the intrinsic content based on a corresponding scene in the
extrinsic content, there is no disclosure or suggestion that the intrinsic content analyzer, the
extrinsic content analyzer and the processor are all included in a single device. Examiner relied
on Fig. 7 to show that the intrinsic/extrinsic content analyzer and the processor are all included in
a single device, as shown; "a command handler (with command API) 702, a properties handler
(with properties API) 704, an event generator (with event API) 706, a cookie manager (with
cookie API) 708, an identifier engine 710, an initialization module 712, a navigator state module
714, a bookmark manager 716" are interconnected to provide the link between intrinsic and
extrinsic content. Specifically the Identifier Engine provides the analysis of both intrinsic and
extrinsic content (paragraph [0302]). Applicant has relied on the same argument stating that
both Courtney and Witteman fails to disclose that the intrinsic content analyzer, the extrinsic

content analyzer and the processor are all included in a single device, however since Lamkin

teaches the limitation as disclosed above the rejections previously presented remains the same in

view of the Examiners arguments presented above.

Claim Rejections - 35 USC § 103

6.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in
> section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are
> such that the subject matter as a whole would have been obvious at the time the invention was made to a person
> having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the
> manner in which the invention was made.

7.   Claims 1-5, 8, 12, 17-23,25, 26, 27, 29 are rejected under 35 U.S.C. 103(a) as being

unpatentable over **Ikezoye et al. (US 6834308 B1)** and in view of **Lampkin et al. (US**

20040220791 A1).

As per claim 1, Ikezoye discloses:

- **A system for integrative analysis of intrinsic,** as (col. 2 lines 51-52) "generates a media

   sample or analytical representation of the media content", **and extrinsic audio-visual**

   data, as (col. 2 lines 53 – 55) "media sample or representation is compared to a database

   of the sampled media content or representations to query and ascertain information

   related to the sample", **the system comprising:**

- **an intrinsic content analyzer, the intrinsic content analyzer being communicatively**

   connected to an audio-visual source, the intrinsic content analyzer being adapted to

   search the audio-visual source for intrinsic data and being adapted to extract

   intrinsic data using an extraction algorithm, as (col. 7 lines 62 – 63), "sampling unit

34 carries out the operation of creating a media sample of the media content played on

the client media player 14", where the media content is the audio-visual source.

- **an extrinsic content analyzer, the extrinsic content analyzer being communicatively**

  connected to an extrinsic information source, the extrinsic content analyzer being

  adapted to search the extrinsic information source and being adapted to retrieve

  extrinsic data using a retrieval algorithm, as (col. 8 lines 26 – 27) "media player 14

  and transmit the sample to the lookup server 12...the lookup server 12 provides the

  information related to the media sample".

- But **Ikezoye** fails to specifically disclose: **a processor configured to correlate the**

  intrinsic data and the extrinsic data for providing a multisource data structure,

  wherein intrinsic analyzer, the extrinsic analyzer an the processor are included in a

  single device

However, **Lampkin** teaches the above limitations as (paragraph [0552]) "the content

management system links the viewer to a corresponding scene (by use of the command

InterActual.SearchTime to go to the specific location within a title) within the DVD-Video… the

text of the screenplay in HTML scrolls with the DVD-Video (e.g., in one of the sub windows) to

give the appearance of being synchronized with the DVD-Video", where the html of the screen

play is the extrinsic data and the specific location within a title) within the DVD-Video is the

intrinsic data and Fig. 7 shows the incorporation of the screen play provided in html, dvd video

and the processor all being incorporated.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of the

invention made to incorporate the teaching of **Lampkin** into the teaching of **Ikezoye** because

one of the ordinary skill in the art would have been motivated to use such a modification for the

purpose of providing related screenplay information based on information related to a video and

audio source. By providing related screenplay information the experience of watching a film is

enhanced by providing more interactive options and information.

As per Claim 2, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein the retrieval of the extrinsic data is based on the extracted intrinsic data**, as

    (col. 8 lines 25 - 30) "media player 14 and transmit the sample to the lookup server

    12...the lookup server 12 provides the information related to the media sample... content-

    related information is received by the user interface 38".

As per Claim 3, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein the extraction and/or retrieval algorithm(s) is/are provided by a module**, as

    (col. 7 lines 62 – 63), "sampling unit 34 carries out the operation of creating a media

    sample of the media content played on the client media player 14", where the sampling

    unit is the module as claimed.

As per Claim 4, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein a query is provided by a user, the query being provided to the extraction**

    algorithm and wherein the intrinsic data is extracted in accordance with the query,

    as (col. 8 lines 20 – 23) "a user my issue a request for content-related information via the

    user-interface 38. This request is communicated to the sampling unit 34 for further

    processing", where the request is claimed query.

As per Claim 5, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein a query is provided by a user, the query being provided to the retrieval algorithm and wherein the extrinsic data is retrieved in accordance with the query**, as (col. 8 lines 20 – 30) "a user my issue a request for content-related information via the user-interface 38. This request is communicated to the sampling unit 34 for further processing... transmit the sample to the lookup server 12...the lookup server 12 provides the information related to the media sample", where the request is claimed query.

As per Claim 8, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein the extrinsic information source is connected to and may be accessed via the Internet (103)**, as (col. 3 lines 23 - 24) "the lookup server is generally connected to the client media players via an Internet connection."

As per Claim 12, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein a feature in a film is analyzed based on information included in the screenplay**, as (col. 8 lines 26 – 27) "media player 14 and transmit the sample to the lookup server 12...the lookup server 12 provides the information related to the media sample" and furthermore (col. 4 lines 50 – 51) "media content source may also be audio CDs, DVD or other formats suitable for presentation on the media play devices" where DVD can contain a film screenplay.

As per Claim 17, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein a high- level information structure (5-9) is generated in accordance with the multi-source data structure**, as (col. 8 lines 30 – 32) "content-related information is received by the user interface 38" where the content-related information returned is high-level as claimed.

As per Claim 18, Claim 17 is incorporated and further Ikezoye discloses:

- **wherein the high- level information structure may be stored on a storage medium**, as
  (col. 8 lines 30 – 32) "content-related information is received by the user interface 38"
  where the user interface is in the client device consisting of a storage medium (col. 6
  lines 30 – 35).

As per Claim 19, Claim 17 is incorporated and further Ikezoye discloses:

- **wherein an update high-level information structure is generated, the updated high-
  level information structure being an already existing high-level information
  structure which is updated in accordance with the multi-source data structure**, as
  (col. 8 lines 30 – 32) "content-related information is received by the user interface 38"
  where the user interface is in the client device consisting of a storage medium (col. 6
  lines 30 – 35) where each time the content-related information is received it is an update.

As per Claim 20, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein the retrieval algorithm is a dynamic retrieval algorithm adapted to
  dynamically update itself by including additional functionalities in accordance with
  retrieved extrinsic data**, as (col. 8 lines 30 – 32) "content-related information is
  received by the user interface 38" where each time the content-related information is
  received it is displayed to the user and therefore stored as an update.

As per Claim 21, Claim 20 is incorporated and further Ikezoye discloses:

- **wherein the additional functionalities is obtained by training the retrieval algorithm
  on a set of features from intrinsic data using labels obtained from the extrinsic data,**
  (col. 9 lines 3 - 8) discloses a log unit that maintains media request such as media, type,

genre or category which assist in training the lookup server by keeping track of what is

requested.

As per Claim 22, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein the training is performed using at least one screenplay**, as (col. 9 lines 3 - 8)

  discloses a log unit that maintains media request such as media, type, genre or category

  which assist in training the lookup server by keeping track of what is requested.

  Furthermore (col. 4 lines 50 – 51) "media content source may also be audio CDs, DVD or

  other formats suitable for presentation on the media play devices" where DVD can

  contain a film screenplay.

As per Claim 23, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein an automatic ground truth identification in a film is obtained based on the**

  multi-source data structure for use in benchmarking algorithms on audio-visual

  content, as (col. 8 lines 47 – 59) "sequentially compares each reference sample in the

  structure 47 to the media sample provided by the media player…The reference that has

  the smallest distance to any frame in the sample is considered a match" which shows the

  multi-source data correlated automatically which is automatic ground truth as claimed.

As per Claim 25, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein an automatic labeling in a film is obtained based on the multi-source data**

  structure, as (col. 8 lines 62 - 63) "This content-related information may include such

  information as song title, artist, and album name".

Claims 26, 27, 28, 29 are method claims for integrative analyses of intrinsic and extrinsic audio-visual source, corresponding to the method claims 1, 17, 20 respectively, and are rejected under the same reason set forth in connection to rejections of claims 1, 17, 20 respectively above.

8.   Claims 10, 11, 24, 28 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Ikezoye and Lampkin** and further in view of **Courtney (**US 5969755 A**).**

As per Claim 10, Claim 1 is incorporated and further Ikezoye and Lampkin fails to disclose:

- **extrinsic data is retrieved based on information extracted from audio video source, wherein the extrinsic content analyzer include knowledge about screenplay grammar**

However, **Courtney** teaches the above limitations as (col. 1 lines 36-42) "consider an on-line movie screenplay (textual data) and a digitized movie (video and audio data). If one were analyzing the screenplay and interested in searching for instances of the word "horse" in the text, many text searching algorithms could be employed to locate every instance of this symbol as desired. Such analysis is common in on-line text databases" which discloses analyzing the screenplay and searching/retrieving the related information from the screenplay from an on-line source based on the video and audio data.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of the invention made to incorporate the teaching of **Courtney** into the teaching of **Ikezoye and Lampkin** because one of the ordinary skill in the art would have been motivated to use such a modification for the purpose of providing related screenplay information based on information extracted from a video and audio source. By providing related screenplay information the

experience of watching a film is enhanced by providing more interactive options and
information.

**As per Claim 11, Claim 1 is incorporated and further Ikezoye and Lampkin does not
disclose:**

- **wherein the identification of persons in a film is obtained by means of the
  screenplay.**

However, **Courtney** teaches the above limitations as (col. 1 lines 36-42) "consider an on-line
movie screenplay (textual data) and a digitized movie (video and audio data). If one were
analyzing the screenplay and interested in searching for instances of the word "horse" in the text,
many text searching algorithms could be employed to locate every instance of this symbol as
desired. Such analysis is common in on-line text databases" which discloses analyzing the
screenplay and searching/retrieving the related information from the screenplay from an on-line
source based on the video and audio data, where related information can a person's identification.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of the
invention made to incorporate the teaching of **Courtney** into the teaching of **Ikezoye and
Lampkin** because one of the ordinary skill in the art would have been motivated to use such a
modification for the purpose of providing related screenplay information based on information
extracted from a video and audio source. By providing related screenplay information the
experience of watching a film is enhanced by providing more interactive options and
information.

**As per Claim 24, Claim 1 is incorporated and further Ikezoye and Lampkin does not
disclose:**

- **wherein an automatic scene content understanding in a film is obtained based on the
  textual description in the screenplay and the audio-visual features from the film
  content**.

However, **Courtney** teaches the above limitations as (col. 1 lines 36-42) "consider an on-line
movie screenplay (textual data) and a digitized movie (video and audio data). If one were
analyzing the screenplay and interested in searching for instances of the word "horse" in the text,
many text searching algorithms could be employed to locate every instance of this symbol as
desired. Such analysis is common in on-line text databases" which discloses analyzing the
screenplay and searching/retrieving the related information from the screenplay from an on-line
source based on the video and audio data.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of the
invention made to incorporate the teaching of **Courtney** into the teaching of **Ikezoye and
Lampkin** because one of the ordinary skill in the art would have been motivated to use such a
modification for the purpose of providing related screenplay information based on information
extracted from a video and audio source. By providing related screenplay information the
experience of watching a film is enhanced by providing more interactive options and
information.

Claim 28 is a method claim for integrative analyses of intrinsic and extrinsic audio-visual
source, corresponding to the method claim 10 and is rejected under the same reason set forth in
connection to rejections of claims 10 respectively above.

9.   Claims 6, 7, 13-16 are rejected under 35 U.S.C. 103(a) as being unpatentable over
Ikezoye and Lampkin and further in view of **Witteman** (US 6243676 B1**).**

As per Claim 6, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein a feature reflected in the intrinsic and extrinsic data include textual, audio and/or visual features,** as (col. 8 lines 60 - 65) discloses the related matching records return to the user include audio and or visual feature.

- But **Ikezoye and Courtney** fails to disclose **textual features.**

  However, **Witteman** teaches the above limitations as (col. 4 line 34) "closed caption text feed is then separated."

  Therefore it would have been obvious to one of the ordinary skill in the art at the time of the invention made to incorporate the teaching of **Witteman** into the teaching of **Ikezoye and Lampkin** because one of the ordinary skill in the art would have been motivated to use such a modification for the purpose of providing additional options and description describing the content so that more relevant information can be retrieved.

As per Claim 7, Claim 1 is incorporated and further Ikezoye discloses:

- **wherein the audio\- visual source is a film (l01)** as (col. 2 lines 53) "media content, such as audio/video played on the media player", where a video played on the media player is a film

- **and wherein the extracted data include textual (104), audio and/or visual features (105, 106)** as (col. 8 lines 60 - 65) discloses the related matching records return to the user include audio and or visual feature.

- But **Ikezoye and Lampkin** fails to disclose **textual features.**

  However, **Witteman** teaches the above limitations as (col. 4 line 34) "closed caption text feed is then separated."

Therefore it would have been obvious to one of the ordinary skill in the art at the time of

the invention made to incorporate the teaching of **Witteman** into the teaching of **Ikezoye and**

Lampkin because one of the ordinary skill in the art would have been motivated to use such a

modification for the purpose of providing additional options and description describing the

content so that more relevant information can be retrieved.

As per Claim 13, Claim 1 is incorporated and further Ikezoye and Lampkin does not

disclose:

- **wherein the correlation of the intrinsic and extrinsic data is time correlation**

  (121), thereby providing a multisource data structure where a feature reflected in

  the intrinsic data is time correlated to a feature reflected in the extrinsic data.

However **Witteman** discloses the above limitation as (**FIG. 3**) which discloses the

linking the extracted text and related searches bases on time.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of

the invention made to incorporate the teaching of **Witteman** into the teaching of **Ikezoye and**

Lampkin because one of the ordinary skill in the art would have been motivated to use such a

modification for the purpose of providing a related content within similar a category of time

which provides a data package consisting of multiple information related to the source.

As per Claim 14, Claim 13 is incorporated and further Ikezoye and Lampkin does not

disclose:

- **wherein the time correlation is obtained by an alignment of a dialogue in the**

  screenplay to the spoken text in the film and thereby providing a timestamped

  transcript of the film.

However **Witteman** discloses the above limitation as (FIG. 3) which discloses the

linking the extracted text and related searches bases on time recognized speech.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of

the invention made to incorporate the teaching of **Witteman** into the teaching of **Ikezoye and**

Lampkin because one of the ordinary skill in the art would have been motivated to use such a

modification for the purpose of providing a related content within similar a category of time

which provides a data package consisting of multiple information related to the source.

As per Claim 15, Claim 14 is incorporated and further Ikezoye and Lampkin does not

disclose:

- **wherein a speaker identification in the film is obtained from the time stamped**
  transcript.

However **Witteman** discloses the above limitation as (col. 4 lines 42 - 43) "process 400

then determines a start of the audio block, indexes the audio block and sends the audio block to

an information store" where determining the start of the audio block is claimed time stamp.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of

the invention made to incorporate the teaching of **Witteman** into the teaching of **Ikezoye and**

Lampkin because one of the ordinary skill in the art would have been motivated to use such a

modification for the purpose of providing a related content within similar a category of time

which provides a data package consisting of multiple information related to the source.

As per Claim 16, Claim 1 is incorporated and further Ikezoye and Lampkin does not

disclose:

- **wherein the screenplay is compared with the spoken text in the film by means of a self-similarity matrix.**

However **Witteman** discloses the above limitation as (FIG. 3) which discloses the linking the extracted text and related searches of the audio based on time and recognized speech.

Therefore it would have been obvious to one of the ordinary skill in the art at the time of the invention made to incorporate the teaching of **Witteman** into the teaching of **Ikezoye and Lampkin** because one of the ordinary skill in the art would have been motivated to use such a modification for the purpose of providing a related content within similar a category of time which provides a data package consisting of multiple information related to the source.

Conclusion

10.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

11.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to DENNIS TRUONG whose telephone number is (571)270-3157.

The examiner can normally be reached on MON - FRI: 7:30 - 5:00 PM EST.

       If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Mahmoudi Tony can be reached on (571) 272-4078. The fax phone number for the

organization where this application or proceeding is assigned is 571-273-8300.

       Information regarding the status of an application may be obtained from the Patent

Application Information Retrieval (PAIR) system. Status information for published applications

may be obtained from either Private PAIR or Public PAIR. Status information for unpublished

applications is available through Private PAIR only. For more information about the PAIR

system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR

system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would

like assistance from a USPTO Customer Service Representative or access to the automated

information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

                                              /Tony Mahmoudi/
                                              Supervisory Patent Examiner, Art Unit
                                              2169

/Dennis Truong/
Examiner, Art Unit 2169